TEXAS A&M UNIVERSITY
PROJECT THEMIS

Technical Report No. 56

DDC
RECEIVED
MAR 7 1978
F

A NEW STATISTICAL APPROACH TO PROJECT SCHEDULING

by

R. L. Sielken, Jr. and H. O. Hartley

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Texas A&M University
Office of Naval Research
Contract N000014-68-A-0140
Project NR047-700

Reproduction in whole or in part
is permitted for any purpose of
the United States Government.

This document has been approved
for public release and sale;
its distribution is unlimited.

ATTACHMENT I

A NEW STATISTICAL APPROACH TO PROJECT SCHEDULING

by

R. L. Sielken, Jr. and H. O. Hartley

THEMIS OPTIMIZATION RESEARCH PROGRAM
Technical Report No.
December 1977

INSTITUTE OF STATISTICS
Texas A&M University

Research conducted through the
Texas A&M Research Foundation
and sponsored by the
Office of Naval Research
Contract N00014-68-A-0140
Project NR047-700

Reproduction in whole or in part
is permitted for any purpose of
the United States Government.

This document has been approved
for public release and sale; its
distribution is unlimited.

ATTACHMENT II

A New Statistical Approach to Project Scheduling

R. L. Sielken, Jr. and H. O. Hartley

Abstract

This paper describes a comprehensive new procedure for determining a minimum cost project schedule when the activities making up the project have durations which are random variables. The cost of an activity is assumed to be a convex piecewise linear function of the activity's mean duration. The objective is to determine the activity mean durations which both minimize the total project cost and insure that the mean of the corresponding project completion time distribution is less than or equal to a specified project deadline. The entire distribution of the project's completion time under the minimum cost schedule is a valuable by-product.

This paper represents a relatively non-mathematical overview of the project scheduling procedure developed in six technical reports prepared under a research contract for the Office of Naval Research. A computer implementation of the procedure may be obtained from Dr. R. L. Sielken, Jr., Institute of Statistics, Texas A&M University, College Station, Texas 77843.

1. The Project Scheduling Problem

A project is composed of several "tasks" or "activities." These activities can be represented by arcs in a directed network. For example, a small project might consist of activities A, B, C, D, and E with the following precedence relationships:

(i) A must be completed before either C or D can be begun;

(ii) B must be completed before D can be begun; and

(iii) C and D must both be completed before E can be begun.

The corresponding network representation is shown in Figure 1. The arc labeled F does not correspond to any "real" activity but is a "dummy" activity merely representing the precedence relation that A must be completed before D can be begun. The points numbered 1, 2, ..., 5 are called nodes. In the network representation of a project the activities originating at a node can be begun only after all activities terminating at that node have been completed.

The time that it actually takes to complete an activity once that activity has been begun is called the activity's duration and is a random variable. The cost of an activity is assumed to be a convex piecewise linear function of the activity's mean duration time. An example of an activity's cost curve is given in Figure 2. In this example TIME(1) is the minimum mean duration time that can be scheduled. TIME(4) is the cheapest mean duration and hence the maximum mean duration that would be scheduled. Of course a linear cost curve is the simplest convex piecewise linear cost function. The more general piecewise behavior, however, frequently arises if there are alternative methods of performing an activity. These methods do not differ in the end result but do differ in the amount of time they take and their cost. For example, to have a mean duration in the interval [TIME(1), TIME(2)] might require the use of a very expensive piece of special equipment while having a mean duration in the interval [TIME(2), TIME(3)] requires only specially trained personnel and having a mean duration in the interval [TIME(3), TIME(4)] just requires varying amounts of standard resources. The form of the activity duration distribution may vary from one time interval to another. For example, the activity duration distribution might be a beta distribution when the mean duration is in [TIME(1), TIME(3)] and approximately a normal distribution when the mean duration is in [TIME(3), TIME(4)].

A project schedule is a specification of each activity's mean duration. The total project cost is simply the sum of the corresponding activity costs. The time to complete the entire project is a random variable whose distribution depends upon the activity duration distributions. The objective is to determine a minimum cost project schedule such that the mean of the corresponding project completion time distribution is less than or equal to a specified project deadline.

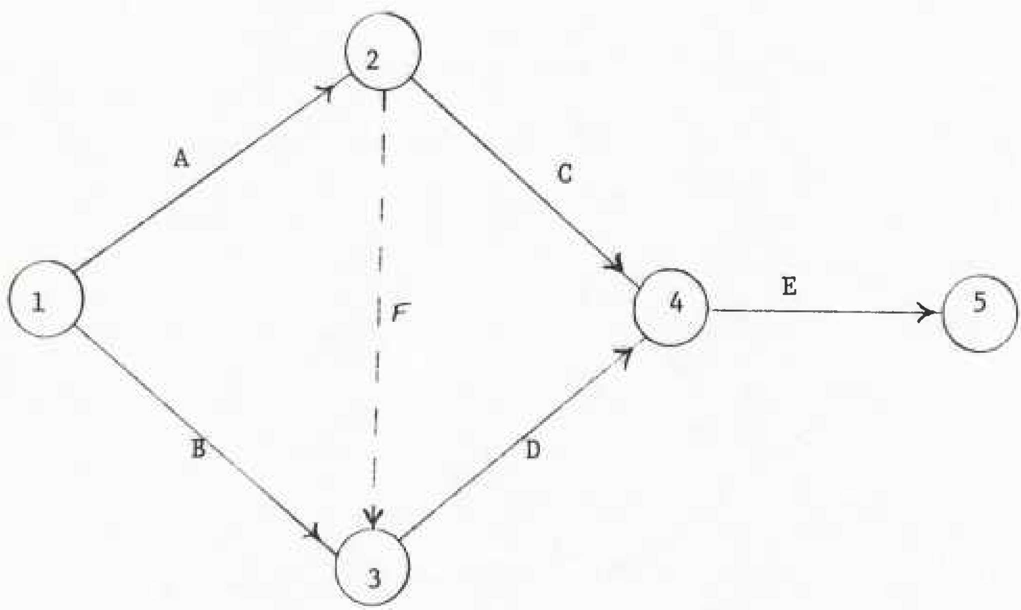


Figure 1. A Small Project Represented as a Directed Network

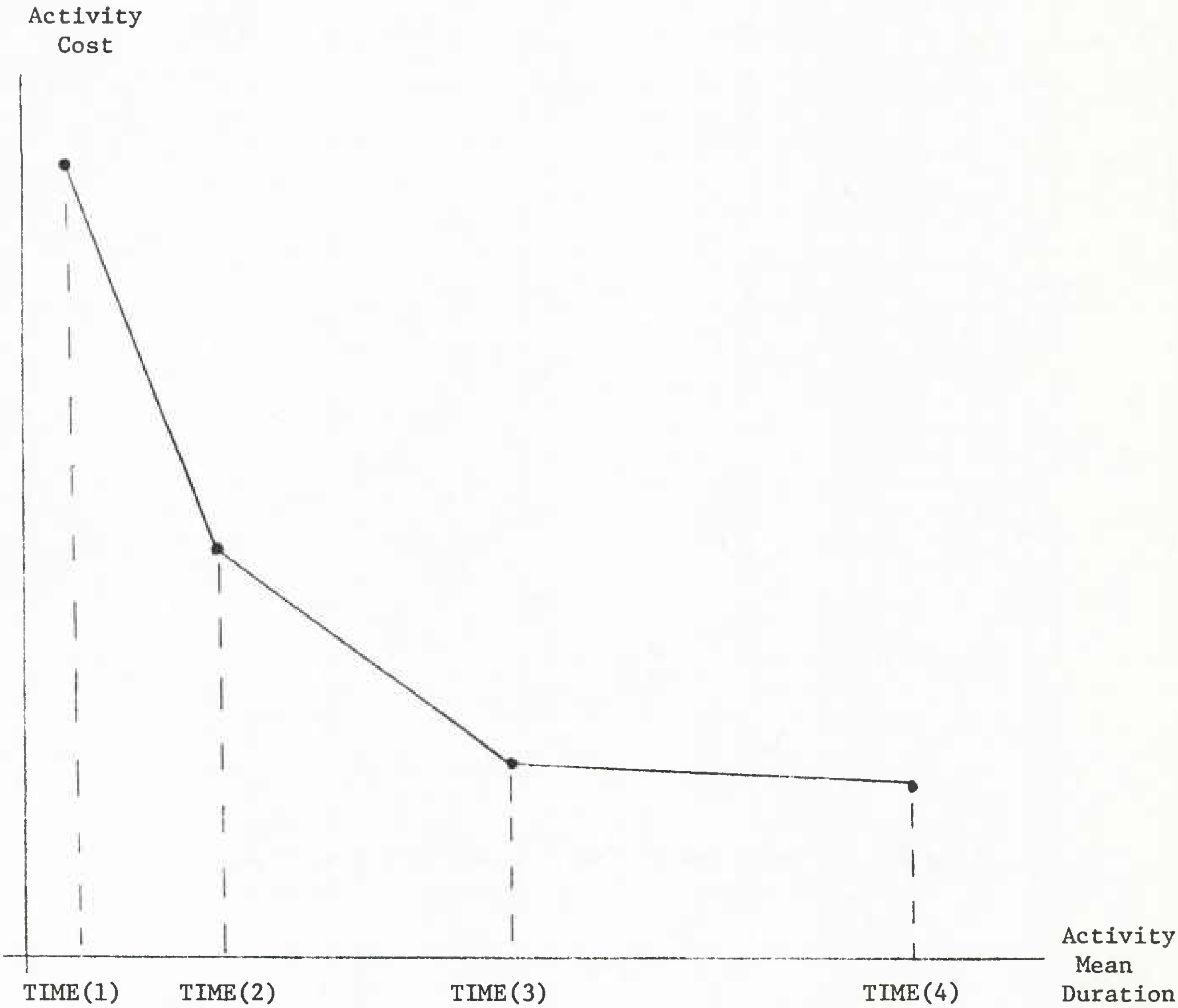


Figure 2. An Activity's Cost as a Function of the Activity's Mean Duration

2. Determining a Project's Completion Time Distribution

Even when each activity's duration distribution is specified, the determination of the project completion time distribution is in general very difficult. The classical approach to the simpler problem of estimating just the mean project completion time is to replace each activity's random duration by its mean duration and then let the corresponding project completion time be the estimate of the project's mean completion time. Unfortunately, this estimate is always less than or equal to the project's mean completion time and can often cause the project manager to greatly underestimate the project's mean completion time. For example, suppose that the project is made up of just n independent activities in parallel, say Y_1, Y_2, ..., Y_n, and that an activity Y has duration 0 with probability $1 - 1/2^m$ and duration 1 with probability $1/2^m$. Then each Y has mean $1/2^m$ and the classical estimate of the project's mean completion time is $1/2^m$. However, since the project's completion time is the maximum of Y_1, Y_2, ..., Y_n, the project's mean completion is really $1 - (1 - 1/2^m)^n$. Thus, if $m = 4$ and $n = 10$, the classical estimate is only 13% of project's mean completion time. Furthermore, as m and n increase the percentage decreases to zero.

In Figure 1 there are really three paths through that network; namely,

$$P_1 = A + C + E,$$

$$P_2 = A + F + D + E,$$

$$P_3 = B + D + E.$$

The project completion time is the maximum of P_1, P_2, and P_3. The classical approach to estimating the project completion time distribution is to determine the maximum path when the random activity durations A, B, ..., F are each replaced by their mean durations and then let the estimate be this particular path's completion time distribution or an approximation thereof. Since this particular path may not be the maximum path for all values of A, B, ..., F, the classical approach overestimates the project completion time distribution. The extent of the overestimation can be illustrated by again considering a project network made up of just n independent activities in parallel. If $G(t)$ is the duration distribution of each activity duration, then the classical estimate of the project completion time distribution would also be $G(t)$ whereas the actual distribution is $[G(t)]^n$. Hence, if there were $n = 4$ activities and the probability that a particular activity's duration is less than or equal to $t = 10$ was 0.8 so that $G(10) = 0.8$, the classical estimate of the probability that the entire project would be completed by $t = 10$ would be 0.8 whereas the actual probability would only be $(0.8)^4 = 0.4096$ - a considerable difference.

From a statistical viewpoint the project completion time is defined easily enough as the maximum of the paths. However, the difficulty is that the paths are not independent since the paths often have activities in common. For example, the paths P_1 and P_2 have activities A and E in common. Theoretical results on behavior of the maximum of dependent random variables are very limited, but research in this area is continuing. Another difficulty is that the number of paths generally increases drastically as the number of activities increases. For example, there are only three paths from the six activities in Figure 1, but there are 69 paths from the 18 activities in Figure 6.

3. Synopsis of a New Project Scheduling Algorithm

In 1974 the development of a new project scheduling procedure was begun with the support of the Office of Naval Research. A detailed documentation of that development is contained in the six technical reports listed in the references. The new project scheduling procedure is an iterative algorithm involving the following five general steps:

Step 1. Deterministic Scheduling: Find a minimum cost project schedule which completes the project by TARGET TIME when each activity's duration is exactly its mean duration and hence deterministic instead of random. (The initial value of TARGET TIME is usually the specified project deadline.)

Step 2. Simplification: Let each activity's duration be a random variable with distribution corresponding to that activity's mean duration chosen during Deterministic Scheduling. Replace various configurations of activities by single activities. The duration distribution for a replacement activity is the distribution of the time to complete all of the activities in the configuration it is replacing. The result of this step is a simplified project network with fewer activities.

Step 3. Decomposition: Partition the simplified project network into several subnetworks in such a way that the resultant subnetworks can be linked together in either series or parallel to form the simplified project network.

Step 4. Subnetwork Analysis: Analyze separately each of the subnetworks determined during Decomposition. Within a subnetwork each activity's duration distribution is approximated by a two-point discrete distribution with matching mean, variance, and third moment. Determine the subnetwork duration distribution corresponding to these discrete activity duration distributions.

Step 5. Synthesis: Combine the approximate subnetwork duration distributions to obtain an approximate completion time distribution for the entire project. If the mean, $\hat{T}$, of this project completion time distribution is sufficiently close to the specified project deadline, the "optimal" project schedule has been found. Otherwise, reset TARGET TIME to

New TARGET TIME = Old TARGET TIME*(Project Deadline/$\hat{T}$)

and return to Step 1.

These five general steps are discussed in more detail in Sections 4, ..., 8 respectively. Section 9 contains a complete example of the iterative algorithm's performance.

4. Deterministic Scheduling

The problem of finding a minimum cost project schedule which conpletes the project by TARGET TIME when each activity's duration is exactly its mean duration can be formulated as a linear programming problem. However, due to the large number of variables and constraints involved, a straightforward linear programming solution would be impractical. Instead the dual of this linear programming problem is considered, further reformulated, and then solved using the very efficient network-flow algorithm described in Dunn and Sielken [1977]. This network-flow algorithm is a generalization of D. R. Fulkerson's algorithm [1961] for solving similar problems with linear activity cost functions. The generalized network-flow algorithm iteratively generates the minimum cost project schedule for every feasible deterministic completion deadline. The corresponding deterministic project cost curve is a convex piecewise linear function of TARGET TIME and a valuable description of the relationship between a project's cost and its deadline. The optimal activity mean durations are linear functions of the TARGET TIME on each linear piece of the project cost curve.

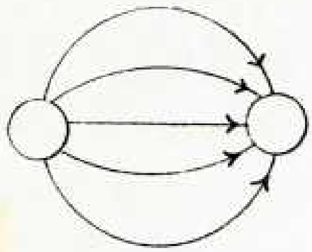
Since the deterministic project scheduler iteratively generates the optimal activity mean durations for all feasible completion deadlines, Step 1 is essentially only performed once. When the general iterative algorithm returns to Step 1 with a new TARGET TIME, finding the optimal activity mean durations is essentially a simple table look-up procedure. For a more complete detailed discussion of Deterministic Scheduling see Dunn and Sielken [1977].

5. Simplification

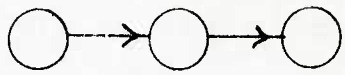
Five configurations of activities for which a single equivalent activity and duration distribution are readily available are depicted in Figure 3. The equivalent single activity duration distributions for the parallel, series, and Wheatstone Bridge configurations were originally identified by Hartley and Wortham [1966] and for the Double Wheatstone Bridge and Criss-Cross configurations by Ringer [1969].

Simplification is an iterative procedure as illustrated in Figure 4. In the special case where Simplification reduces the project network down to just one activity as in Figure 4, the project completion time distribution is directly determined so that Steps 3 and 4 are skipped. Although a reduction to one activity is a very special case, reductions of over 50% are quite common.

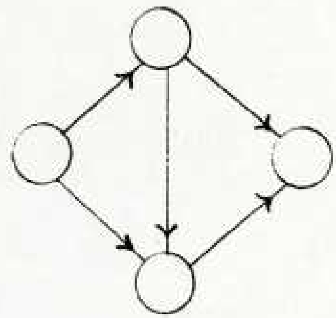
The structure of the simplified network is identified the first time Step 2 is performed. In subsequent iterations only the distribution of the activity durations in the simplified network change.



Activities in
Parallel

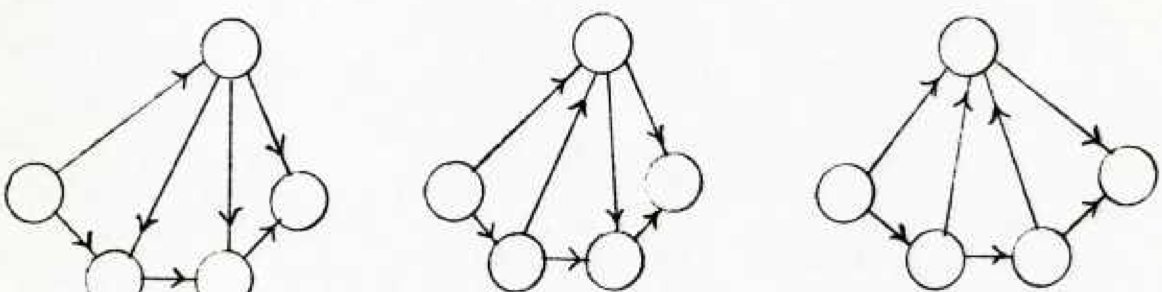


Two Activities in
Series



Wheatstone Bridge

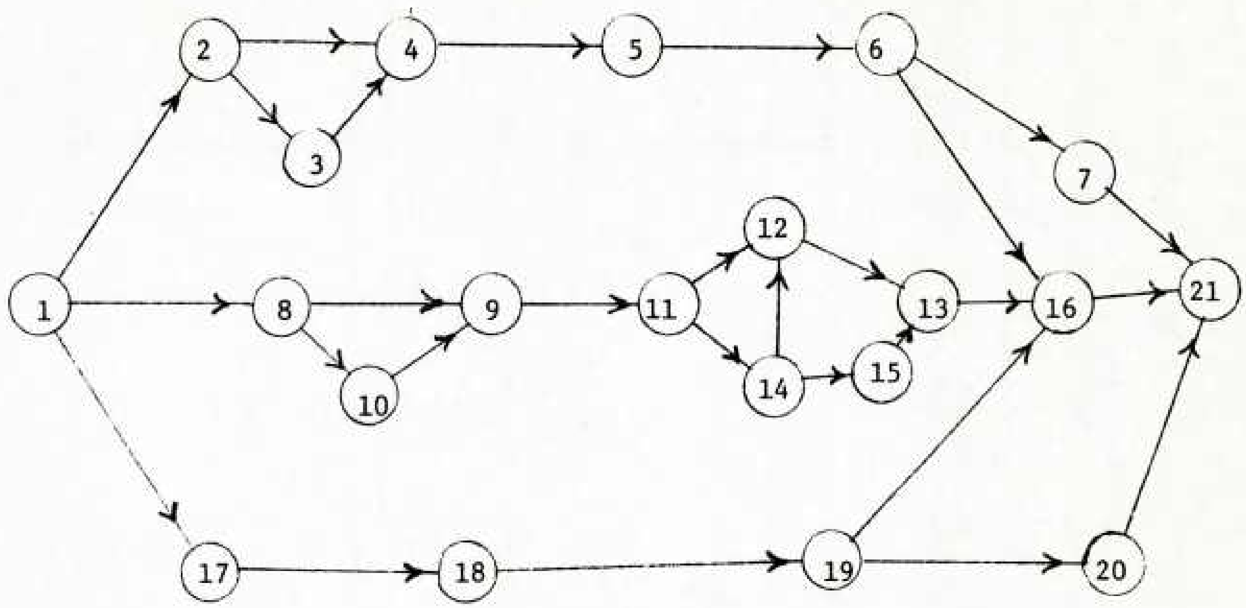
Double Wheatstone Bridges



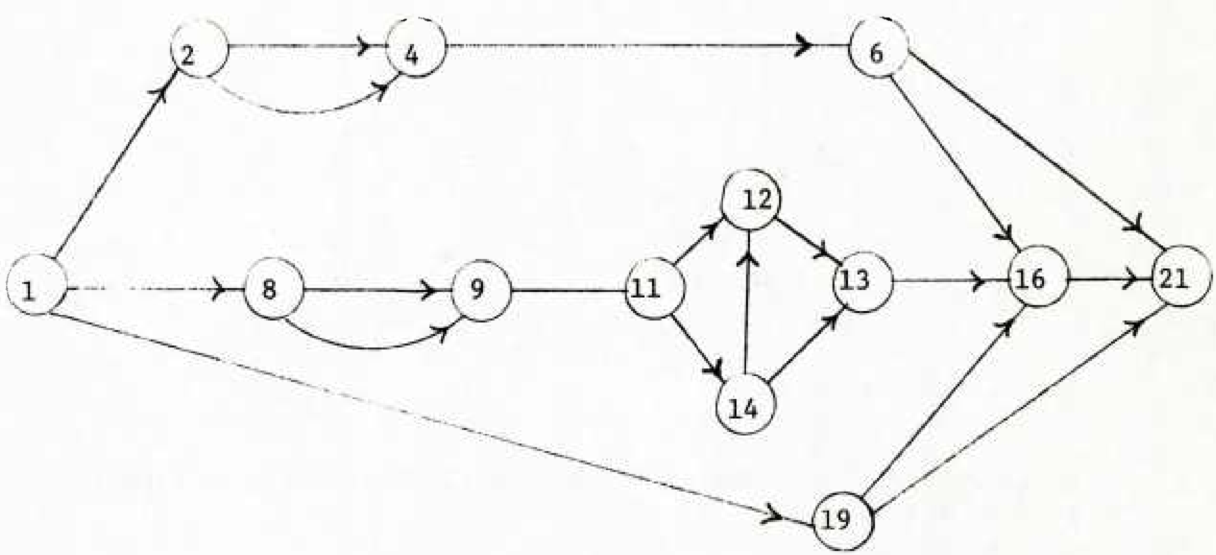
Criss - Crosses

Figure 3. Activity Configurations Which Can be Readily Replaced by a Single Equivalent Activity

Figure 4. The Iterative Simplification of a Project Network

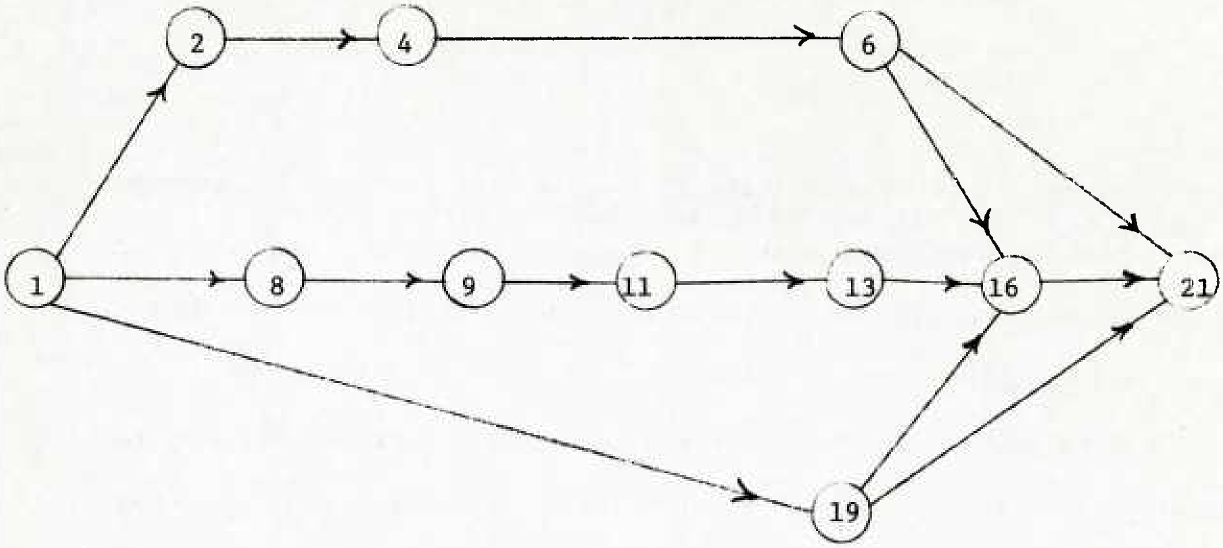


Original Project Network

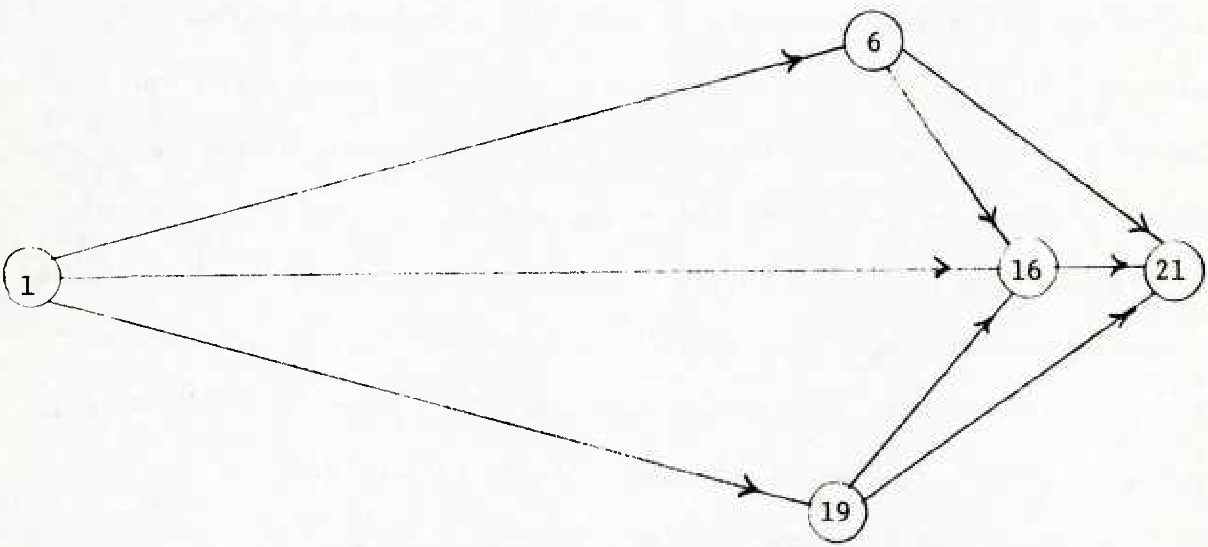


Project Network After Activities in Series
Replaced by Equivalent Single Activities

Figure 4. (Continued)



Project Network After Activities in Parallel and Activities in a Wheatstone Bridge Replaced by Equivalent Single Activities



Project Network After Activities in Series Replaced by Equivalent Single Activities

Simplified Project Network After Activities in a Double Wheatstone Bridge Replaced by an Equivalent Single Activity

6. Decomposition

The objective of Decomposition is to partition the simplified project network into the simplest possible subnetworks subject to the constraint that the subnetwork duration distributions can be easily combined to yield the project completion time distribution.

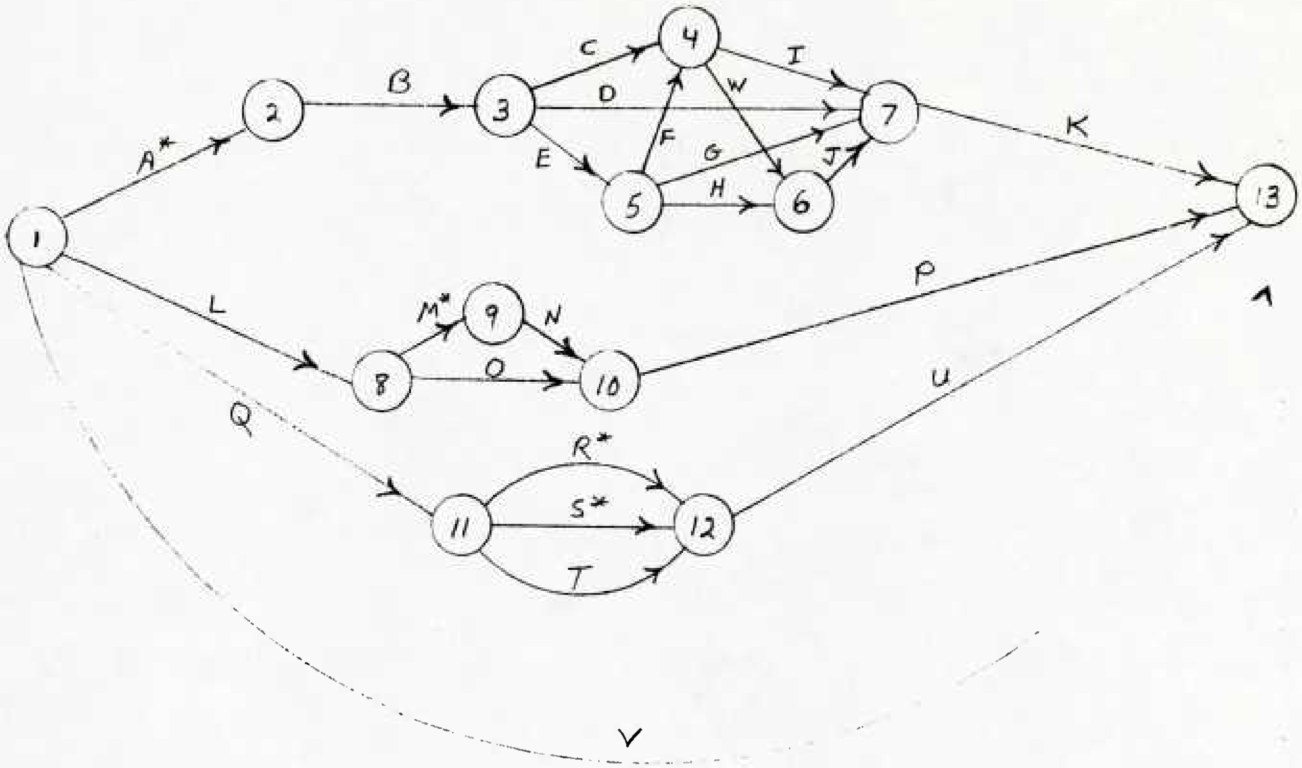
The simplified project network can be suitably partitioned using the following iterative procedure. First the simplified network is searched for subnetworks that begin at the beginning of the simplified network, end at the end of the simplified network, and are in parallel. Each parallel subnetwork is then subdivided into a sequence of smaller subnetworks that are in series. Each series subnetwork is then searched for parallel subnetworks. The partitioning into parallel and series subnetworks continues until no subnetwork can be further partitioned. This iterative partitioning procedure is illustrated in Figure 5.

As in Simplification this partitioning of the simplified project network into subnetworks could have also included subnetwork configurations of the Wheatstone Bridge, Double Wheatstone Bridge, and Criss-Cross forms; however, the apparent frequency of these subnetwork configurations does not seem to justify the additional programming effort.

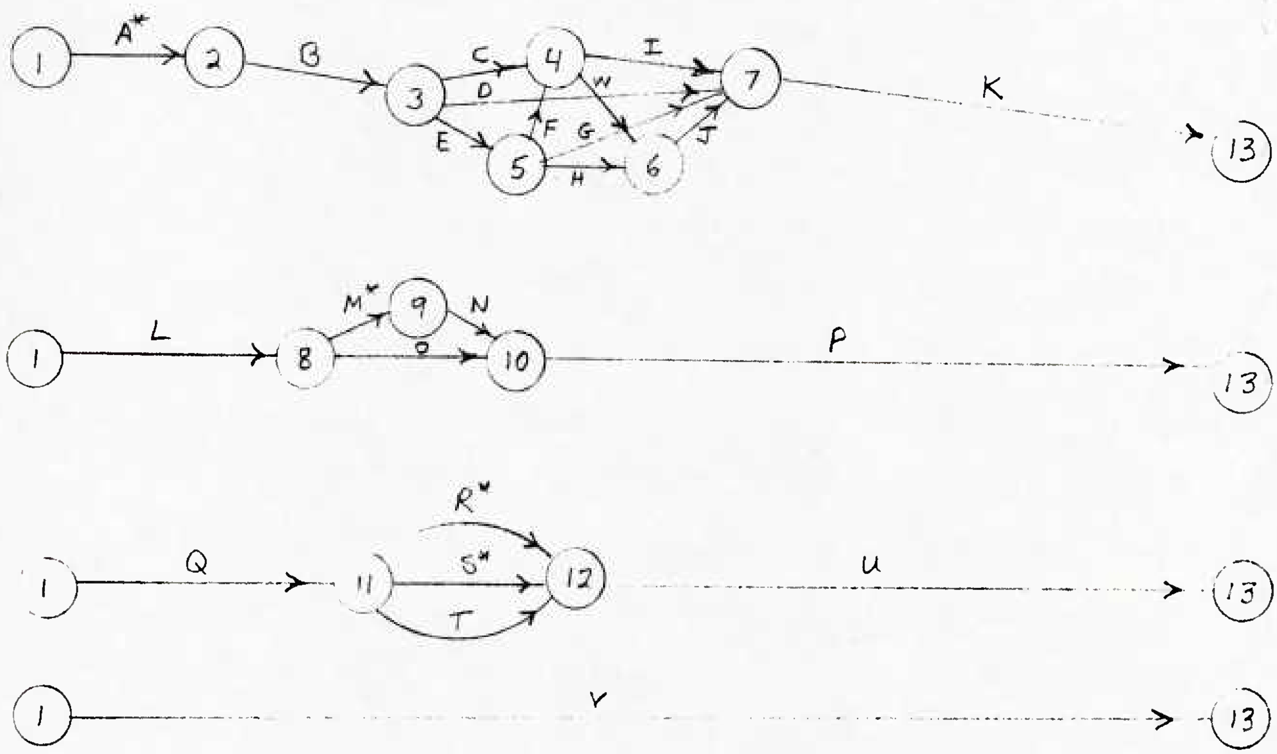
Since Steps 1 and 2 do not change the structure of the simplified project network and the partitioning of that network does not depend on the activity duration distributions, Decomposition is only done once and is really skipped when the general iterative algorithm returns to Step 3.

Decomposition is documented in Sielken and Fisher [1976].

Figure 5. The Decomposition of a Simplified Project Network*



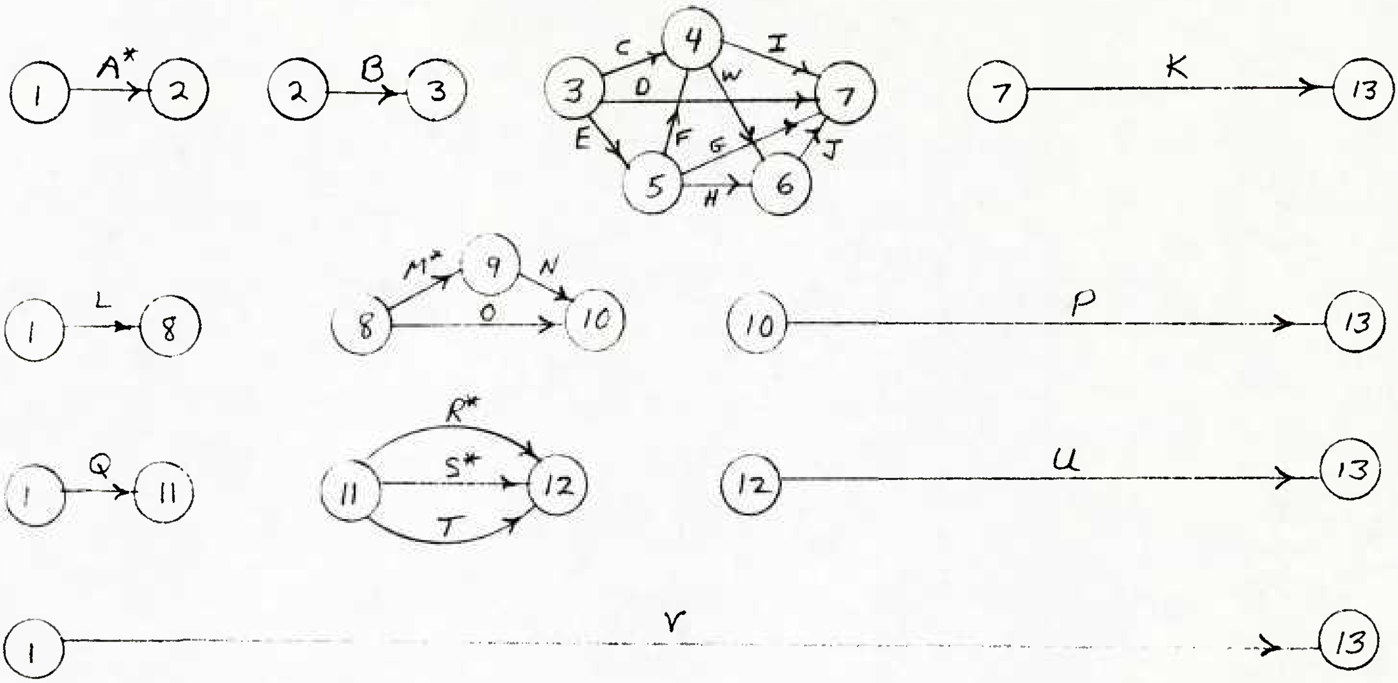
Simplified Project Network



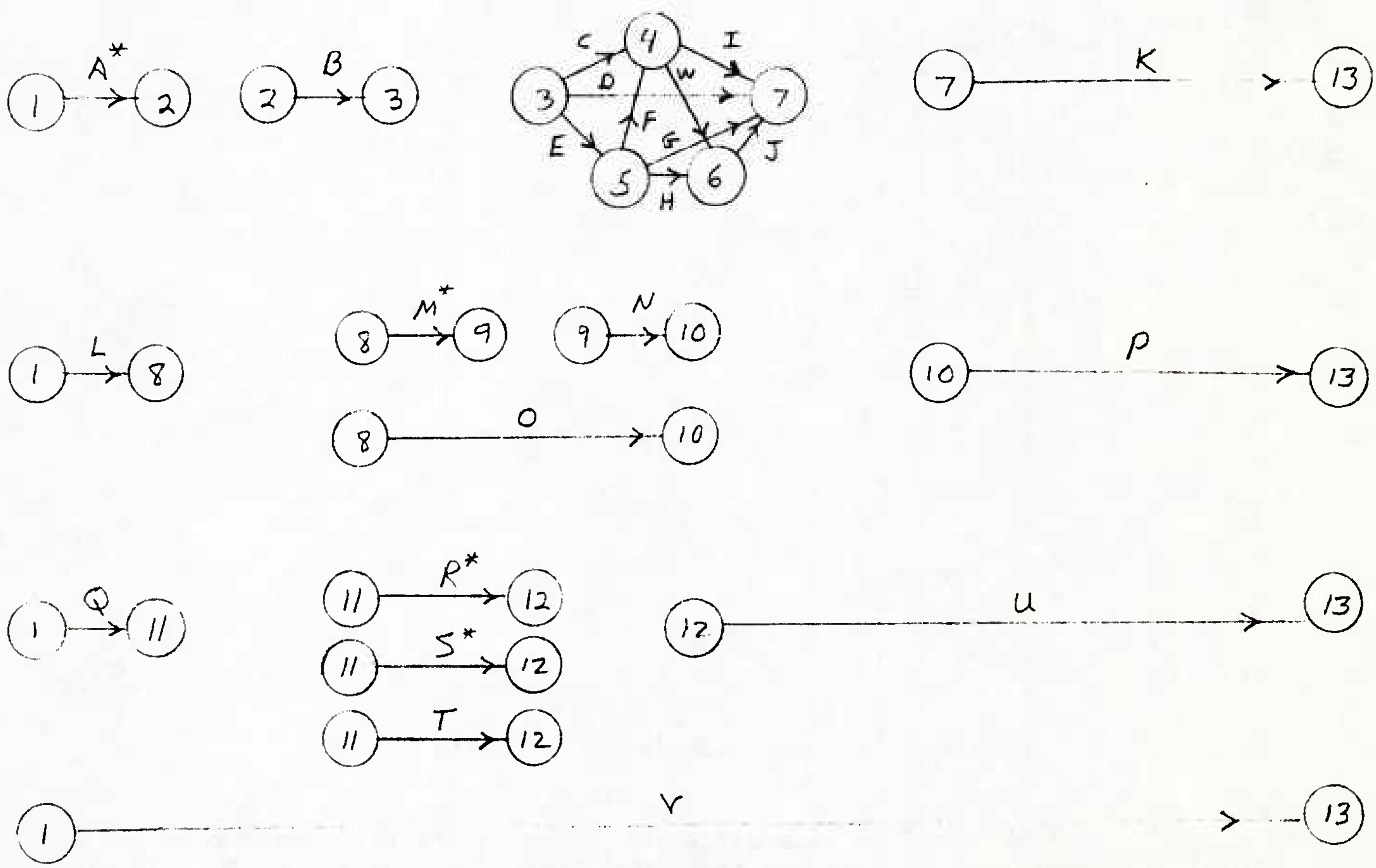
Partitioning into Parallel Subnetworks

The arcs A, M*, R*, and S* each represent non-simplifiable subnetworks and not single activities. Otherwise this "Simplified Project Network" could be further simplified.

Figure 5. (Continued)



Partitioning into Series Subnetworks



Final Partitioned Simplified Project Network

7. Subnetwork Analysis

In Step 4, Subnetwork Analysis, the objective is to approximate each subnetwork's duration distribution. At the end of Step 2 each activity in the subnetwork has a specified duration distribution. This distribution is now approximated by a two-point discrete distribution. In particular, the activity is now conceptualized as having two possible duration times, say L for a lower duration and U for an upper duration. The probability that the activity duration is L is assumed to be p, and correspondingly the probability that the activity duration is U is assumed to be 1 - p. The values of L, U, and p are chosen so that the mean, variance, and third moment of the discrete distribution are the same as the mean, variance, and third moment of the activity's specified duration distribution. The subnetwork's exact duration distribution is approximated by the subnetwork duration distribution when each activity has its two-point discrete duration distribution.

If each activity had a non-random duration, the subnetwork duration would be the duration of the longest path through the subnetwork. The longest path is commonly referred to as the critical path and the activities on that path are called critical activities.

To determine the subnetwork's approximate duration distribution, each activity duration is temporarily set equal to its mean duration, $m = pL + (1 - p)U$, and the critical path determined. The activities on this original critical path are certainly some of the most important activities in determining the subnetwork duration distribution. Hence these original critical activities are the first activities put into a set called IMPORTANT.

Some originally non-critical activities could become critical if the duration of an originally critical activity were decreased. Hence one originally critical activity has its duration decreased to $m - \lambda s$ where s is the activity's standard deviation, $s = [pL^2 + (1 - p)U^2 - m^2]^{\frac{1}{2}}$, and λ is a positive algorithm parameter. The new critical path is determined and each new critical activity not in IMPORTANT is added to IMPORTANT. This procedure is repeated for every originally critical activity.

An originally non-critical activity could become critical if its actual duration was greater than its mean duration. Hence one originally non-critical activity has its duration set equal to $m + \theta s$ where θ is an algorithm parameter while all other activities are assigned their mean durations. The new critical path is determined and each new critical activity not in IMPORTANT is added to IMPORTANT. This procedure is repeated for every originally non-critical activity.

After the composition of IMPORTANT has been determined, a lower bound on the subnetwork's approximate duration distribution can be obtained by considering the activity durations for non-IMPORTANT activities at their U values and the activity durations for IMPORTANT activities at all possible combinations of their L and U values. An upper bound on the subnetwork's approximate duration distribution can be determined by considering the activity durations for non-IMPORTANT activities at their L values and the activity durations for IMPORTANT activities at all possible combinations of their L and U values.

Although the composition of IMPORTANT depends upon both λ and θ, the most important parameter is θ. In particular, for any fixed λ, as θ increases the lower bound on the subnetwork's approximate duration distribution increases monotonically while the upper bound decreases monotonically. In fact, for any fixed λ, there is a finite value of θ such that the lower bound and upper bound are exactly equal to the subnetwork's approximate duration distribution. Since the computational effort in determining the upper and lower bounds increases as θ increases, the most practical procedure is actually to determine the upper and lower bounds for a few θ values and then extrapolate to ascertain the subnetwork's approximate duration distribution.

Computational experience with the Subnetwork Analysis procedure for approximating a subnetwork duration distribution has been highly satisfactory and a substantial improvement over classical approaches. At first glance, a Monte Carlo approach would appear to be equally satisfactory. In fact, it would have been easier to prepare a computer implementation of the Monte Carlo approach than the Subnetwork Analysis procedure. However, the Monte Carlo approach is less practical in the sense of computer execution time (cost) and theoretically less appealing since all activities are given equal consideration in the Monte Carlo approach whereas Subnetwork Analysis concentrates more heavily on the activities most likely to influence the subnetwork duration.

The Subnetwork Analysis procedure was first developed in Arseven, Hartley, Ringer, and Sielken [1974] and then improved in Ringer, Sielken and Spoeri [1976]. The many details and proofs omitted in the above discussion are given in these two references.

8. Synthesis

The project's approximate completion time distribution can be determined by combining the approximate subnetwork duration distributions. When the project network is decomposed in Step 3, the result is a network of subnetworks with any two connected subnetworks being either in series or in parallel. Let SUB_1, and SUB_2 be any two such subnetworks, and let the corresponding approximate subnetwork duration distributions be $\hat{F}_1$ and $\hat{F}_2$. If SUB_1 and SUB_2 are in series, then the approximate duration distribution for SUB_1 and SUB_2 combined is

$$\hat{F}(t) = \sum_{s \leq t} \hat{F}_2(t - s)\hat{f}_1(s) = \sum_{s \leq t} \hat{F}_1(t - s)\hat{f}_2(s) \tag{8.1}$$

where $\hat{f}_1$ and $\hat{f}_2$ are the discrete probability density functions corresponding to $\hat{F}_1$ and $\hat{F}_2$ respectively. If SUB_1 and SUB_2 are in parallel, then the approximate duration distribution for SUB_1 and SUB_2 combined is

$$\hat{F}(t) = \hat{F}_1(t) \cdot \hat{F}_2(t). \tag{8.2}$$

By repeatly combining subnetworks that are connected either in series or in parallel, the project's approximate completion time distribution is obtained.

Once the project's approximate completion time distribution has been determined, the project's approximate mean completion time, $\hat{T}$, can be calculated and compared with the project deadline. If the project manager feels that $\hat{T}$ is sufficiently close to the project deadline, say within 5%, then the project schedule just determined in Step 1 is considered "optimal." Otherwise, a new project schedule must be determined by returning to Step 1 with a new TARGET TIME.

Step 1 only requires that the project would be completed by TARGET TIME if each activity's duration was exactly its mean duration. On the other hand, $\hat{T}$ takes into consideration the random nature of an activity's duration and hence will generally exceed TARGET TIME. The difficulty is in deciding how much less than the project deadline should TARGET TIME be in order that the corresponding $\hat{T}$ be sufficiently close to the project deadline. The algorithm iteratively updates its estimate of this TARGET TIME by

$$\text{New TARGET TIME} = \text{Old TARGET TIME}*(\text{Project Deadline}/\hat{T}).$$

The initial TARGET TIME would usually be the project deadline but could be chosen somewhat less than the project deadline. A typical sequence of TARGET TIMEs and $\hat{T}$'s for a project deadline of 400 might be

TARGET TIME = 400, $\hat{T}$ = 500,

TARGET TIME = 320, $\hat{T}$ = 360,

TARGET TIME = 356, $\hat{T}$ = 408.

Since the algorithm approximates the project's entire completion time distribution for each tentative schedule determined in Step 1, the quantity $\hat{T}$ in the above discussion could just as easily be a specified percentile of the project completion time distribution. For example, the project manager might wish a minimum cost project schedule such that the probability of the project being completed before the project deadline is .90. In this case, the new TARGET TIMEs would be determined with $\hat{T}$ being the 90th percentile of the project's approximate completion time distribution instead of the project's approximate mean completion time.

Of course, each time Steps 1-5 are performed a new $\hat{T}$, approximate project completion time distribution, and project cost are generated. This sequence of $\hat{T}$'s and project costs supplements the deterministic project cost curve in describing the impact of the project deadline on the project cost.

9. An Example of the Project Scheduling Algorithm's Performance

A small project network is depicted in Figure 6. The relationship between each activity's mean duration and its cost is given in Table 1. The project scheduling algorithm also requires that the activity's duration distribution be specified at the midpoint of each time interval on the convex piecewise linear cost function, i.e., when the activity's mean duration is [TIME(1) + TIME(2)]/2, [TIME(2) + TIME(3)]/2, etc. This information is also given in Table 1. The algorithm assumes that if an activity's mean duration is not at the midpoint but at c times the midpoint and still in the same time interval, then the activity's duration distribution has the same form (Normal, Beta, Constant, etc.) but with a new variance equal to c^2 times the variance at the midpoint. Thus, for example, if activity A's mean duration is 28, its cost is 34, and activity A's duration distribution is Beta [10,40] with mean 28 and variance $(28/25)^2 36$.

With a project deadline of 110 the algorithm's iterative determination of the minimum cost project schedule is as follows:

Step 1. Deterministic Scheduling: The shortest feasible project completion time when each activity duration is its mean duration is found to be 90 by determining the longest path through the project network when each activity's duration is equal to its minimum mean duration, TIME(1). Similarly, the longest such feasible project completion time is found to be 135 by determining the longest path through the project network when each activity's duration is equal to its maximum mean duration. The minimum cost schedule which complete the project by TARGET TIME when each activity's duration is its mean duration is determined for each value of TARGET TIME between 90 and 135. The corresponding optimal activity

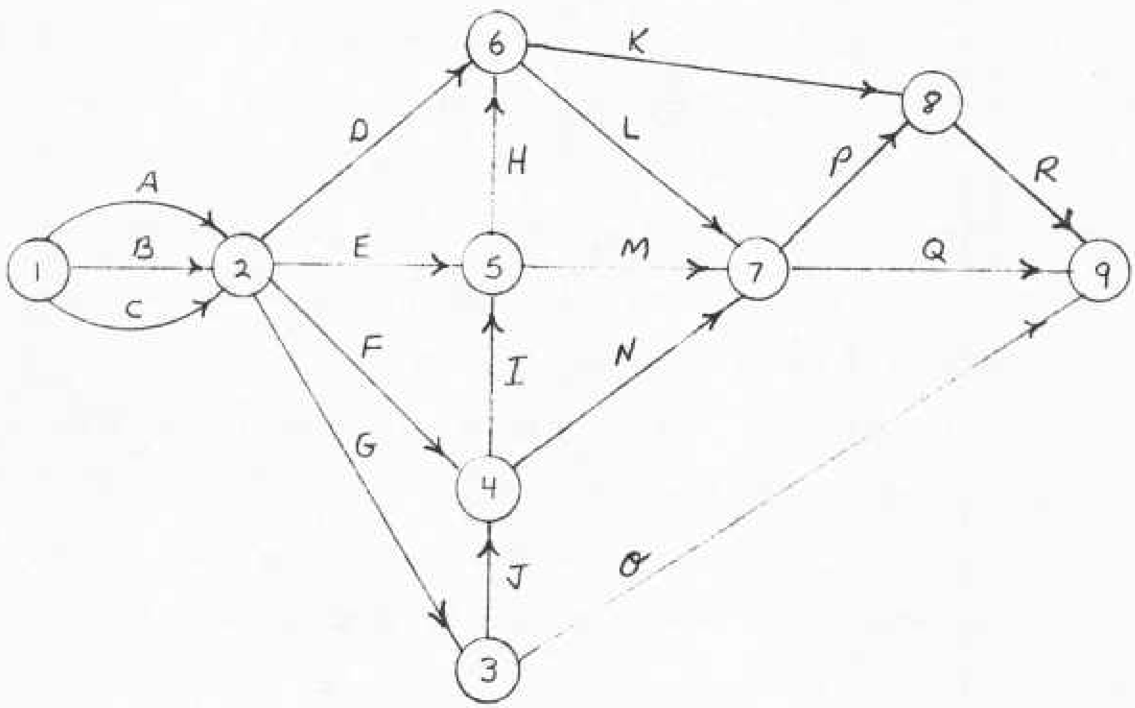


Figure 6. Example Project Network

Table 1. Activity Costs and Duration Distributions

Activity	Mean Durations	Cost	Duration Distribution at Midpoint
A	TIME(1) = 10	100	Beta on [5,20] with mean 12.5 and variance 16
	TIME(2) = 15	70	Normal with mean 17.5 and variance 25
	TIME(3) = 20	50	Beta on [10,40] with mean 25 and variance 36
	TIME(4) = 30	30	
B	TIME(1) = 8	60	Normal with mean 11 and variance 9
	TIME(2) = 14	45	
C	TIME(1) = 10	50	Constant Duration = 10
D	TIME(1) = 50	200	Normal with mean 60 and variance 100
	TIME(2) = 70	160	
E	TIME(1) = 32	45	Normal with mean 36 and variance 18
	TIME(2) = 40	40	
F	TIME(1) = 20	64	Beta on [10,50] with mean 26 and variance 70
	TIME(2) = 32	48	
G	TIME(1) = 13	30	Normal with mean 16 and variance 7
	TIME(2) = 19	25	
H	TIME(1) = 20	60	Normal with mean 25 and variance 12
	TIME(2) = 30	50	Normal with mean 49 and variance 9
	TIME(3) = 35	48	
I	TIME(1) = 5	60	Constant Duration = 5
J	TIME(1) = 18	62	Beta on [10,40] with mean 22 and variance 6
	TIME(2) = 26	49	
K	TIME(1) = 10	40	Constant Duration = 10
L	TIME(1) = 6	75	Normal with mean 8 and variance 5
	TIME(2) = 10	50	

Table 1. (Continued)

Activity	Mean Durations	Cost	Duration Distribution at Midpoint
M	TIME(1) = 36	225	Normal with mean 40 and variance 80
	TIME(2) = 44	175	
N	TIME(1) = 30	400	Normal with mean 35 and variance 25
	TIME(2) = 40	300	
O	TIME(1) = 60	250	Normal with mean 70 and variance 144
	TIME(2) = 80	210	
P	TIME(1) = 4	52	Constant Duration = 4
Q	TIME(1) = 2	100	Beta on [1,6] with mean 3 and variance 1
	TIME(2) = 4	90	
R	TIME(1) = 4	110	Normal with mean 5 and variance 2
	TIME(2) = 6	80	

mean durations are given in Table 2. The project cost curve is depicted in Figure 7.

The initial activity mean durations are those corresponding to TARGET TIME = 110; namely,

A = 17,	F = 32,	K = 10,	P = 4,
B = 14,	G = 13,	L = 10,	Q = 4,
C = 10,	H = 34,	M = 44,	R = 6.
D = 70,	I = 5,	N = 40,	
E = 39,	J = 21,	O = 80,	

Step 2. Simplification: The only configuration of activities which can be readily replaced by an equivalent single activity is A, B, and C in parallel. If the replacement activity is denoted by ABC, then the simplified project network consists of the single activity ABC and the original activities D, E, ..., R. The initial duration distribution for ABC as a function of t is

$$F_{ABC}(t) = F_A(t)F_B(t)F_C(t)$$

where $F_A(t)$ denotes a normal distribution with mean 17 and variance $(17/17.5)^2 25$, $F_B(t)$ denotes a normal distribution with mean 14 and variance $(14/11)^2 9$, and $F_C(t)$ denotes the distribution for a constant duration of 10.

Step 3. Decomposition: The simplified project network is partitioned into two subnetworks in series. The first subnetwork, SUB_1, consists of the single activity ABC. The second subnetwork, SUB_2, consists of the activities D, E, ..., R.

Table 2. Optimal Activity Mean Durations for All Feasible TARGET TIME's

TARGET TIME	135-D	134-D	128-D	124-D	114-D	113-D	108-D	106-D	105-D	104-D	100-D	96-D	92-D
Range of D	$0 \le D \le 1$	$0 \le D \le 6$	$0 \le D \le 4$	$0 \le D \le 10$	$0 \le D \le 1$	$0 \le D \le 5$	$0 \le D \le 2$	$0 \le D \le 1$	$0 \le D \le 1$	$0 \le D \le 4$	$0 \le D \le 4$	$0 \le D \le 4$	$0 \le D \le 2$
Project Cost When D=0	1552.0	1552.4	1557.4	1563.9	1583.9	1586.1	1606.1	1614.6	1620.2	1626.2	1660.2	1702.8	1747.8
Activity													
A	30	30	30	30-D	20	20-D	15	15	15-D	14-D	10	10	10
B	14	14	14	14	14	14	14	14	14	14-D	10	10	10
C	10	10	10	10	10	10	10	10	10	10	10	10	10
D	70	70	70	70	70	70	70	70	70	70	70-D	66-D	62
E	40	40	40	40	40-D	39	39-D	37-D	36	36	36	36	36
F	32	32	32	32	32	32	32	32-D	31	31	31	31	31
G	19	19-D	13	13	13	13	13	13	13	13	13	13	13
H	35-D	34	34	34	34	34	34	34	34	34	34-D	30-D	26
I	5	5	5	5	5	5	5	5	5	5	5	5	5
J	26	26	26-D	22	22-D	21	21-D	19-D	18	18	18	18	18
K	10	10	10	10	10	10	10	10	10	10	10	10	10
L	10	10	10	10	10	10	10	10	10	10	10	10	10
M	44	44	44	44	44	44	44	44	44	44	44-D	40-D	36
N	40	40	40	40	40	40	40	40	40	40	40	40	40
O	80	80	80	80	80	80	80-D	78-D	77	77	77-D	73-D	69-D
P	4	4	4	4	4	4	4	4	4	4	4	4	4
Q	4	4	4	4	4	4	4	4	4	4	4	4	4
R	6	6	6	6	6	6	6	6	6	6	6	6	6-D

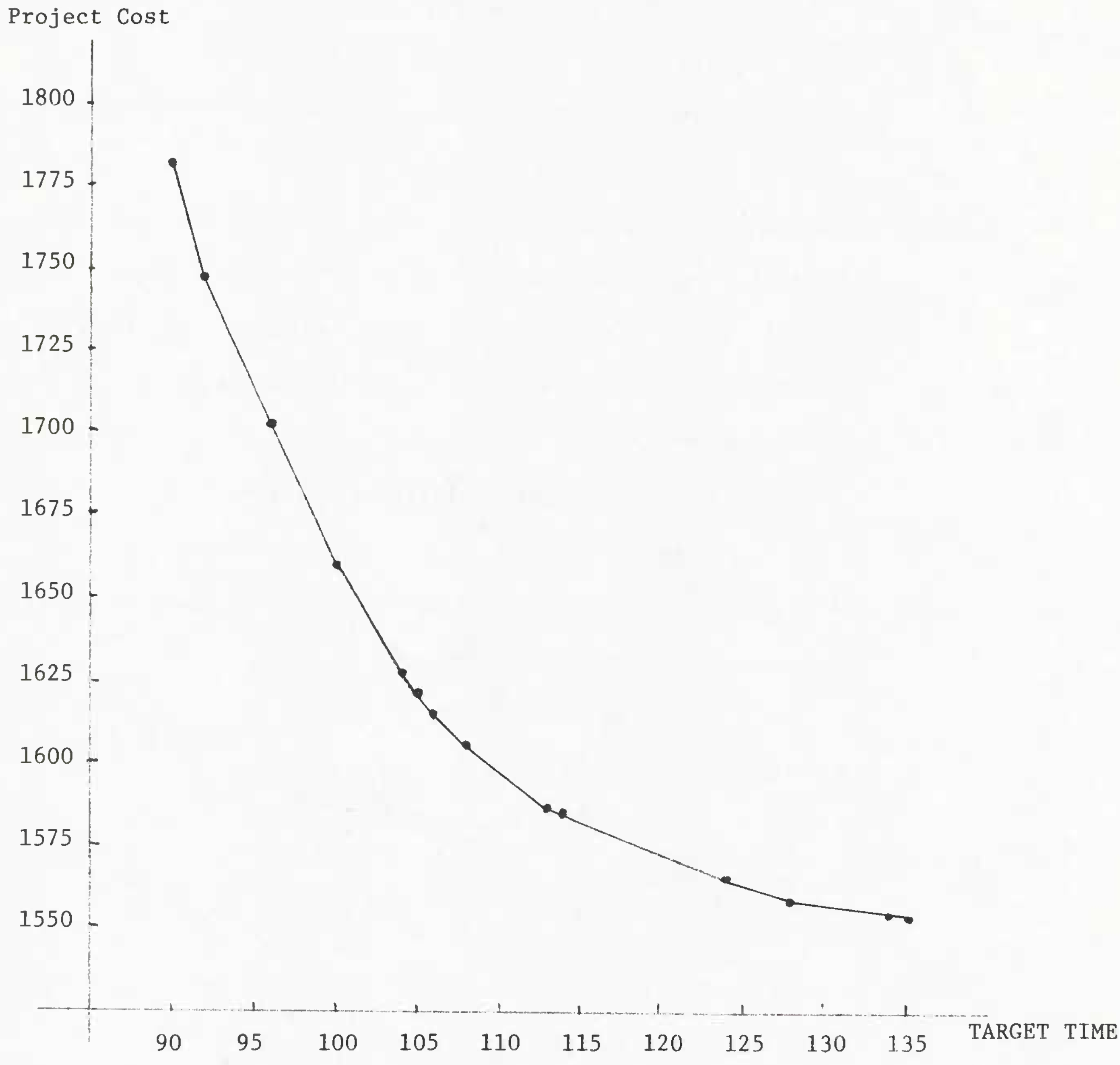


Figure 7. Total Project Cost as a Function of TARGET TIME

Step 4. Subnetwork Analysis: The activity duration distributions for activities ABC,D, ..., R are approximated by two-point discrete distributions with matching mean, variance, and third moment. Then the subnetwork duration distribution corresponding to these discrete activity duration distributions is determined for SUB_1 and SUB_2 respectively.

Step 5. Synthesis: Since SUB_1 and SUB_2 are in series, their approximate duration distributions determined in Step 4 are combined using equation (8.1) to yield an approximate project completion time distribution. The mean of this distribution is 131.77, so that the new TARGET TIME is

$$\text{New TARGET TIME} = 110(110/131.77) = 91.83$$

Return to Step 1.

Step 1. Deterministic Scheduling: Using Table 2 the optimal activity mean durations for TARGET TIME = 91.83 are

A = 10,	F = 31,	K = 10,	P = 4,
B = 10,	G = 13,	L = 10,	Q = 4,
C = 10,	H = 26,	M = 36,	R = 5.83.
D = 62,	I = 5,	N = 40,	
E = 36,	J = 18,	O = 68.83,	

Step 2. Simplification: The structure of the simplified project network never changes after the first iteration. Only the duration distribution of ABC needs to be redetermined.

Step 3. Decomposition: The partitioning of the simplified project network never changes after the first iteration.

Step 4. Subnetwork Analysis: The new activity duration distributions are approximated by new two-point discrete distributions. Then new approximate duration distributions for SUB_1 and SUB_2 are determined.

Step 5. Synthesis: The new approximate project mean completion time is 107.27, so that the new TARGET TIME is

$$\text{New TARGET TIME} = 91.83(110/107.27) = 94.16.$$

Return to Step 1.

On the fifth iteration the project scheduling algorithm finds that the optimal activity mean durations for TARGET TIME = 93.88 have a corresponding project completion time distribution with mean 109.37. If 109.37 was sufficiently close to the specified project deadline of 110, then these optimal activity mean durations would constitute the minimum cost project schedule.

10. Monitoring a Partially Completed Project

Once the project has been scheduled and work begins, the project manager continually gains new information about activity durations which can be used to update the project's estimated completion time distribution. Suppose, for example, that after so many days have elapsed, activities A, B, C, and G in Figure 6 have been completed, activities D, E, and F have been in progress for 23 days, and activities J and O have been in progress for seven days. Then an updated estimate of the project completion time distribution can be obtained by repeating Steps 2-5 once with

(i) The random durations for activities A, B, C, and G replaced by the observed durations;

(ii) The duration distribution for each partially completed activity replaced by its conditional duration distribution given that it has been worked on for an observed number of days; and

(iii) The duration distributions for activities which have not been begun still equal to those determined in the minimum cost project schedule.

Furthermore, if after completing Step 5 and examining the updated estimate of the project completion time distribution, the project manager could reschedule the uncompleted portion of the project by letting the project scheduling algorithm return to Step 1 and proceed again from there.

If a substantial portion of the project has already been completed, the project network could be restated as in Figure 8 before returning to the steps in the project scheduling algorithm. Such a restatement would not effect the updated project completion time distribution or any rescheduling but would be solely for the purpose of reducing the algorithm's computational effort and operating cost. The dashed activities in Figure 8 represent the partially completed activities. These activities would not have their conditional duration distributions altered during any rescheduling if the project manager was reluctant to change the way an activity was being performed in the midst of that performance.

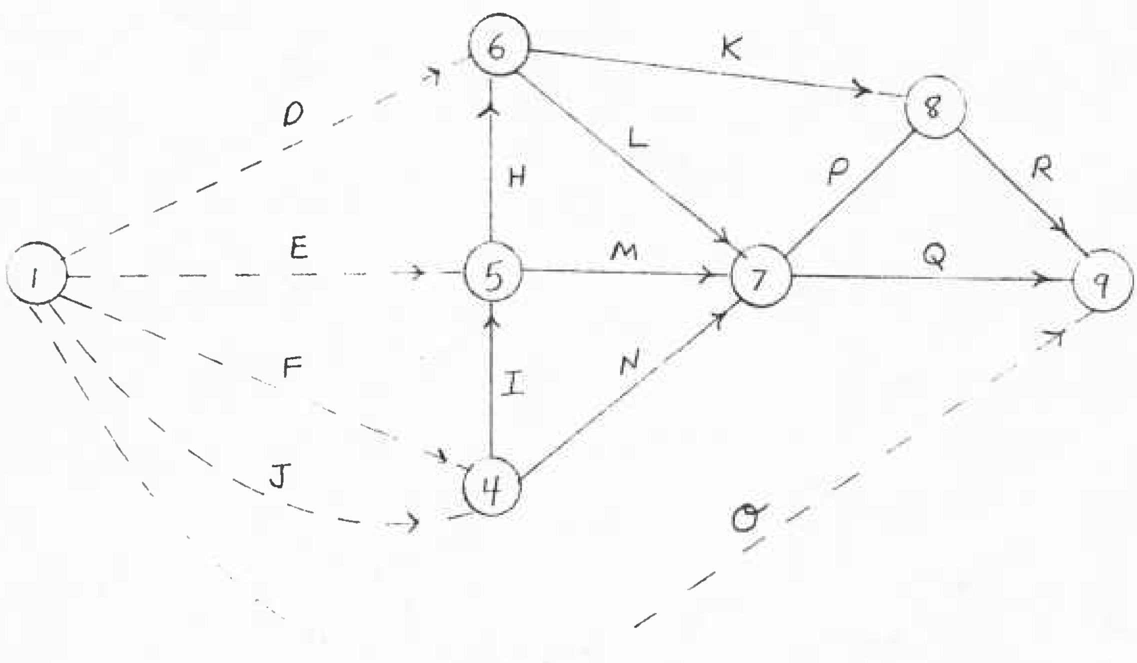


Figure 8. Restatement of a Partially Completed Project.

11. Concluding Remarks

The new project scheduling procedure allows the project scheduler to specify

(i) the precedences among the project's activities,

(ii) the relationship between an activity's cost and its mean duration,

(iii) the manner in which an activity's actual duration varies about its mean duration, and

(iv) a deadline for either the project's mean completion time or a prescribed percentile of the project completion time distribution.

In return the project scheduler receives

(i) a minimum cost project schedule which delineates each activity's mean duration time,

(ii) an estimate of the distribution of the project completion time,

(iii) information on the trade-off between the project's minimum cost and its specified deadline, and

(iv) a tool for monitoring the project's progress and, if need be, rescheduling.

An exciting feature of this new project scheduling procedure is that it simultaneously incorporates the desire to minimize the project cost and the realization that an activity's duration is not necessarily a fixed quantity exactly equal to its prescribed duration but rather a random quantity varying about a prescribed duration. No longer must the project scheduler either (i) choose a reasonable cost schedule which heuristically hedges against the randomness in the activities he quesses will be critical, or (ii) choose a reasonable schedule which should probably finish before the deadline and then guess where he can save money without disturbing the suspected completion time too much. By considering both cost and randomness together in one systematic algorithm, the new project scheduling procedure eliminates this guesswork.

The authors wish to acknowledge their gratitude to the Office of Naval Research for the support of this research under contracts N00014-68-A-0140 and N00014-76-C-0038. Several present and past members of the Institute of Statistics at Texas A&M University have also contributed to the development of this new project scheduling procedure and its computer implementation: E. Arseven (Lederle Laboratories), P. P. Biemer, T. C. Baker, Jr., C. S. Dunn, N. E. Fisher (Compucon Inc.), L. J. Ringer, and R. K. Spoeri (Bureau of the Census).

References

Fulkerson, D. R., "A Network Flow Computation for Project Cost Curves," Management Science, Vol. 7, No. 2, (January, 1961), pp. 167-178.

Hartley, H. O. and A. W. Wortham, "A Statistical Theory for PERT Critical Path Analysis," Management Science, Vol. 12, No. 10, (June, 1966), pp. 469-481.

Ringer, L. J., "Numerical Operators for Statistical PERT Critical Path Analysis," Management Science, Vol. 16, No. 2, (October, 1969), pp. B-136 - B-143.

Technical Reports

Biemer, Paul P. and R. L. Sielken, Jr., "Incorporating Project Cost Considerations into Stochastic PERT," THEMIS Project Technical Report No. 52, November, 1975.

Dunn, C. S. and R. L. Sielken, Jr., "Statistical PERT: An Improved Project Scheduling Algorithm," THEMIS Project Technical Report No. 55, February, 1977.

Sielken, R. L., Jr. and Norman E. Fisher, "Statistical PERT: Decomposing a Project Network," THEMIS Project Technical Report No. 50, January, 1976.

Sielken, R. L., Jr., H. O. Hartley and R. K. Spoeri, "Statistical PERT: An Improved Subnetwork Analysis Procedure," THEMIS Project Technical Report No. 51, January, 1976.

Sielken, R. L., Jr., L. J. Ringer, H. O. Hartley and E. Arseven, "Statistical Critical Path Analysis in Acyclic Stochastic Networks: Statistical PERT," THEMIS Project Technical Report No. 48, November, 1974.

Spoeri, R. K., L. J. Ringer and R. L. Sielken, Jr., "A Statistical Procedure for Optimization of PERT Network Scheduling Systems," THEMIS Project Technical Report No. 53, April, 1976.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Texas A&M University

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP
Unclassifed

3. REPORT TITLE
A NEW STATISTICAL APPROACH TO PROJECT SCHEDULING

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

5. AUTHOR(S) (First name, middle initial, last name)
R. L. Sielken Jr. and H. O. Hartley

6. REPORT DATE
December 15, 1977

7a. TOTAL NO. OF PAGES

7b. NO. OF REFS
3

8a. CONTRACT OR GRANT NO.
N00014-68-A-0140

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
Number 56

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distributuion is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Office of Naval Research

13. ABSTRACT

See page iv of report.

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Project scheduling						
Minimum cost						
Random activity durations						
Project completion time distribution						
Statistical PERT						
Critical path methodology						
Networks						

BASIC DISTRIBUTION LIST FOR
UNCLASSIFIED TECHNICAL REPORTS

	Copies
Operations Research Office of Naval Research (Code 434) Arlington, Virginia 22217	3
Director, Naval Research Laboratory Attn: Library, Code 2029 (ONRL) Washington, D.C. 20390	2
Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12
Defense Logistics Studies Information Exchange Army Logistics Management Center Fort Lee, Virginia 23801	1
Technical Information Center Naval Research Laboratory Washington, D.C. 20390	6
Office of Naval Research New York Area Office 207 West 24th Street New York, New York 10011 Attn: Dr. J. Laderman	1
Director, Office of Naval Research Branch Office 1030 East Green Street Pasadena, California 91101 Attn: Dr. A.R. Laufer	1
Lt. Dennis Johnston - ED3 Manned Spacecraft Center NASA Houston, Texas 77058	1
Mr. William Dejka Code 360 Naval Electronics Laboratory Center San Diego, California 92132	1
Professor Lucien A. Schmit, Jr. Room 6731, Boelter Hall School of Engrng. & Applied Sci. University of California Los Angeles, California 90024	1
Director Office of Naval Research Branch Office 495 Summer Street Boston, Massachusetts 02210 Attn: Dr. A.L. Powell	1
Director Office of Naval Research Branch Office 219 South Dearborn Street Chicago, Illinois 60604 Attn: Dr. A.R. Dawe	1
Office of Naval Research San Francisco Area Office 1076 Mission Street San Francisco, California 94103	1
Technical Library Naval Ordnance Station Indian Head, Maryland 20640	1
Bureau of Naval Personnel Navy Department Technical Library Washington, D.C. 20370 STOP 82	2
Library, Code 0212 Naval Postgraduate School Monterey, California 93940	1
Naval Ordnance Station Louisville, Kentucky 40214	1
Library Naval Electronics Laboratory Center San Diego, California 92152	1
Naval Ship Engineering Center Philadelphia Division Technical Library Philadelphia, Pennsylvania 19112	1
Dr. A.L. Slafkosky Scientific Advisor Commandant Marine Corps (Code AX) Washington, D.C. 20380	1

	Copies
Naval Applied Science Laboratory Technical Library, Code 222 Flushing & Washington Avenues Brooklyn, New York 11251	1
Naval Undersea Warfare Center 3202 E. Foothill Boulevard Pasadena, California 91107 Technical Library	1
University of Chicago Statistical Research Center Chicago, Illinois 60637 Attn: Prof. W. Kruskal	1
Stanford University Department of Statistics Stanford, California 94305 Attn: Prof. G.J. Lieberman	1
Florida State University Department of Statistics Tallahassee, Florida 32306 Attn: Prof. I.R. Savage	1
Princeton University Department of Mathematics Princeton, New Jersey 08540 Attn: Prof. J.W. Tukey	1
Stanford University Department of Statistics Stanford, California 94305 Attn: Prof. T.W. Anderson	1
University of California Department of Statistics Berkeley, California 94720 Attn: Prof. P.J. Bickel	1
University of Washington Department of Mathematics Seattle, Washington 98105 Attn: Prof. Z.W. Birnbaum	1
Cornell University Department of Mathematics White Hall Ithaca, New York 14850 Attn: Prof J. Kiefer	1

	Copies
Harvard University Department of Statistics Cambridge, Massachusetts 02139 Attn: Prof. W.G. Cochran	1
Columbia University Department of Industrial Engineering New York, New York 10027 Attn: Prof. C. Derman	1
New York University Institute of Mathematical Sciences New York, New York 10453 Attn: Prof. W.M. Hirsch	1
University of North Carolina Statistics Department Chapel Hill, North Carolina 27515 Attn: Prof. W.L. Smith and Prof. M.R. Leadbetter	1
Purdue University Division of Mathematical Sciences Lafayette, Indiana 47079 Attn: Prof. H. Rubin	1
University of California, San Diego Department of Mathematics P.O. Box 109 La Jolla, California 92038 Attn: Prof. M. Rosenblatt	1
Florida State University Department of Statistics Tallahassee, Florida 32306 Attn: Prof. R.A. Bradley	1
New York University Department of Industrial Engineering & Operations Research Bronx, New York 10453 Attn: Prof. J.H.K. Kao	1
University of Wisconsin Department of Statistics Madison, Wisconsin 53706 Attn: Prof. G.E.P. Box	1

Logistics Research Project	Copies
The George Washington University 707 - 22nd Street, N.W. Washington, D. C. 20037 Attn: Dr. W. H. Marlow	1
International Business Machines Corporation P.O. Box 218, Lamb Estate Yorktown Heights, New York 10598 Attn: Dr. Alan Hoffman	1
University of California Management Sciences Research Project Los Angeles, California 90024 Attn: Dr. J.R. Jackson	1
Harvard University Department of Economics Cambridge, Massachusetts 02138 Attn: Prof. K.J. Arrow	1
Cowles Commission for Research in Economics Yale University New Haven, Connecticut 06520 Attn: Prof. Martin Shubik	1
Carnegie-Mellon University Graduate School of Industrial Administration Pittsburgh, Pennsylvania 15213 Attn: Prof. G. Thompson	1
University of California Department of Economics Berkeley, California 94720 Attn: Prof. R. Radner	1
University of California Operations Research Center Institute of Engineering Research Berkeley, California 94720 Attn: Prof. D. Gale	1
University of California Graduate School of Business Administration Los Angeles, California 90024 Attn: Prof. J. Marschak	1

The University of Michigan	Copies
Department of Mathematics, W.E. Ann Arbor, Michigan 48104 Attn: Prof. R.M. Thrall	1
Princeton University Department of Mathematics Princeton, New Jersey 08540 Attn: Prof. A.W. Tucker	1
Case Western Reserve University Systems Research Center Cleveland, Ohio 44106 Attn: Prof. M. Mesarovic	1
University of Texas Department of Mathematics Austin, Texas 78712 Attn: Dr. A. Charnes	1
Stanford University Department of Operations Research Stanford, California 94305 Attn: Dr. D.L. Iglehart	1
University of Delaware Department of Mathematics Newark, Delaware 19711 Attn: Prof. R. Remage, Jr.	1
Stanford University Department of Operations Research Stanford, California 94305 Attn: Prof. F.S. Hillier	1
Dr. Claude-Alain Burdet Asst. Prof. Industrial Admin. Carnegie-Mellon University Pittsburgh, Pennsylvania 15213	1
Stanford University Department of Operations Research Stanford, California 94305 Attn: Prof. G. B. Dantzig	1
Chief of Naval Research (Code 436) Department of the Navy Arlington, Va. 22217	1
Science Librarian Kresge Library Oakland University Rochester, Michigan 48063	1

	Copies
Purdue University Graduate School of Ind. Admin. Lafayette, Indiana 47907 Attn: Prof. Andrew Whinston	1
Princeton University Department of Economics and Sociology Princeton, New Jersey 08540 Attn: Prof. O. Morgenstern	1
Decision Studies Group 845 Page Mill Road Palo Alto, California 94304 Attn: Dr. Stuart Bessler	1
University of California Department of Engineering Los Angeles, California 90024 Attn: Prof. R.R. O'Neill	1
Maritime Transportation Research Board National Academy of Sciences 2101 Constitution Avenue Washington, D.C. 20418 Attn: RADM J.B. Oren USCG (RET)	1
Stanford University Department of Operations Research Stanford, California 94305 Attn: Prof. A.F. Veinott, Jr.	1
Texas A&M Foundation College Station, Texas 77843 Attn: Prof. H.O. Hartley	1
Case Western Reserve University Cleveland, Ohio 44106 Attn: Prof. B.V. Dean	1
University of California Center for Research in Management Science Berkeley, California 94720 Attn: Prof. W.I. Zangwill	1
U.S. Naval Postgraduate School Department of Operations Research and Economics Monterey, California 93940 Attn: Prof. C.R. Jones	

	Copies
Captain Allen D. Dayton AFOSR (SRMA) 1400 Wilson Blvd. Arlington, Virginia 22209	1
Dr. Paul Murrill Project Themis Department of Chemical Engineering Louisiana State University Baton Rouge, Louisiana	1
Director Army Materials & Mechanics Research Center Attn: Mr. J. Bluhm Watertown, Massachusetts 02172	1
Commanding Officer U.S. Army Ballistic Research Laboratories Attn: Dr. John H. Giese Aberdeen Proving Ground, Maryland 21005	1
University of Iowa Department of Mechanics and Hydraulics Iowa City, Iowa 52240 Attn: Prof. W.F. Ames	1
Office of Naval Research Resident Representative 1105 Guadalupe Lowich Building Austin, Texas 78701 Attn: Mr. Francis M. Lucas	1
Dr. Jerome Bracken Institute of Defense Analyses 400 Army-Navy Drive Arlington, Virginia 22202	1
Professor Richard L. Fox School of Engineering Case Western Reserve University Cleveland, Ohio 44106	1
Director, National Security Agency Ft. George G. Meade, Maryland 20755 Attn: Dr. Joseph Blum, R44	1

University of Connecticut Department of Statistics Storrs, Connecticut 06268 Attn: Prof. H.O. Posten	Copies 1
Prof. Emanuel Parzen Statistics Department State Univ. New York at Buffalo 4230 Ridge Lea Road Amherst, New York 14226	 1
ARCON Corporation Lakeside Office Park North Avenue at Route 128 Wakefield, Massachusetts 01880 Attn: Dr. A. Albert	 1
Stanford University Department of Statistics Stanford, California 94305 Attn: Prof. H. Chernoff	 1
Yale University Department of Statistics New Haven, Connecticut 06520 Attn: Prof. L.A. Savage	 1
Rutgers-The State University Statistics Center New Brunswick, New Jersey 08903 Attn: Prof. H.F. Dodge	 1
Yale University Department of Statistics New Haven, Connecticut 06520 Attn: Prof. F. J. Anscombe	 1
Purdue University Division of Mathematical Sciences Lafayette, Indiana 47907 Attn: Prof. S.S. Gupta	 1
Cornell University Department of Industrial Engineering Ithaca, New York 19850 Attn: Prof. R.E. Bechhofer	 1
Dr. Jay Lewallen E-D-13 Theory and Analysis Division NASA Manned Spacecraft Center Houston, Texas 77058	 1
Professor Geoffrey S. Watson, Chairman Department of Statistics Princeton University Fine Hall Princeton, N. J. 08540	Copies 1
Stanford University Department of Mathematics Stanford, California 94305 Attn: Prof. S. Karlin	 1
University of Sheffield Department of Probability and Statistics Sheffield 10, ENGLAND Attn: Prof. J. Gani	 1
University of California Operations Research Center Institute of Engineering Research Berkeley, California 94720 Attn: Prof. R.E. Barlow	 1
Stanford University Department of Statistics Stanford, California 94305 Attn: Prof. H. Solomon	 1
Applied Mathematics Laboratory Naval Ships Research Development Center Washington, D.C. 20007	 1
Systems Analysis Division Room BE760, Pentagon Washington, D.C. 20350 Attn: Mr. A.S. Rhodes, Op-964	 1
Department of Statistics University of North Carolina Chapel Hill, North Carolina 27515 Attn: Prof. M.R Leadbetter	 1
Southern Methodist University Department of Statistics Dallas, Texas 75222 Attn: Prof. D.B. Owen	 1
Israel Institute of Technology Technion Haifa, ISRAEL Attn: Prof. P. Naor	

	Copies
Office of Gifts & Exchanges Library Texas A&M University College Station, Texas 77843	2
Dr. Del Cilmer Naval Weapons Center (Code 607) China Lake, California 93555	1

58043